

PENGELOMPOKAN KABUPATEN/KOTA DI JAWA TIMUR BERDASARKAN INDIKATOR PENDIDIKAN TAHUN 2013 MENGGUNAKAN ANALISIS HIERARCHIAL CLUSTER

Elok Fitriani Rafikasari

Fakultas Ekonomi dan Bisnis Islam (FEBI), IAIN Tulungagung

E-mail: elokfitriani@ymail.com

Abstract

National development efforts in improving quality of education especially in East Java are to create and to apply the policy that in accordance with education level in each district. Making policy is aimed to achieve awareness and capability to contend especially in education field. In the context of assisting the manufacture of the policy needs to be done a grouping of the education level of community in the district in East Java to understand the extent of the development of the spread of the level of schooling. The purpose of this study is to find groups of districts in East Java have characteristic together in the field of education. While the method of analysis is that are used hierarchical clusters. From this analysis obtained two groups where the first group a number of 29 the districts which have the level of education higher than the second group with 9 of districts.

Kata Kunci: *Clustering, Level Pendidikan, Analisis Hierarchical Cluster*

PENDAHULUAN

Sumber daya manusia yang berkualitas merupakan parameter dari kesuksesan pembangunan nasional. Suatu negara dikatakan maju dan berkembang bukan dilihat dari kekayaan alamnya yang melimpah

tetapi dilihat dari seberapa banyak sumber daya manusia yang baik dan berkualitas yang dimilikinya. Ujung tombak dalam kemajuan bangsa adalah pendidikan karena pendidikan yang berkualitas akan menghasilkan sumber daya manusia yang berkualitas pula. Dunia pendidikan berfungsi untuk memproduksi sumber daya manusia yang berkualitas dalam berbagai bidang. Tenaga yang berkualitas diharapkan mampu menjadi dinamisor dalam pembangunan. Hal ini dikarenakan sumber daya manusia yang berkualitas menjadi penentu gerak dan laju pembangunan.

Permasalahan yang sering muncul adalah isu tentang rendahnya mutu pendidikan dalam berbagai jenjang di Indonesia. Hal ini diperkuat dengan data *Education for All Global Monitoring Report* tahun 2005 yang menyebutkan bahwa populasi buta huruf di Indonesia berkisar 18,4 juta orang¹. Selain itu berdasarkan laporan capaian program *Education for All* (EFA) *Global Monitoring Report* 2008 menyatakan bahwa EFA *Development Index* (EDI) Indonesia berada di peringkat 71 untuk tingkat buta huruf² yang mengisyaratkan bahwa masih lemahnya manajemen pendidikan di Indonesia yang harus segera mendapatkan penanganan untuk dicari solusi terbaiknya. Untuk melakukan penanganan yang tepat perlu dilakukan pengelompokan tingkat pendidikan. Pengelompokan dengan metode *Hierarchical Cluster* dilakukan untuk mengetahui sejauh mana perkembangan penyebaran tingkat pendidikan masyarakat dalam rangka membantu pembuatan kebijakan.

TINJAUAN PUSTAKA

Analisis Komponen Utama (*Principal Component Analysis*)

Analisis komponen utama atau *principal component analysis* adalah ilmu statistika dalam kelompok analisis multivariat yang digunakan untuk mengelompokkan variabel-variabel yang memiliki korelasi linier

¹Ade Sahputra, "Peran UNESCO dalam Pemberantasan Buta Aksara di Indonesia tahun 2007-2012", (*Jom FISIP*, Volume 1 No. 2-Oktober 2014), 2

² Davy Hendri, 2008, *Kaji Ulang Indikator Pembangunan Pendidikan*, <http://targetmdgs.org>

signifikan menjadi sejumlah komponen utama yang dapat mewakili keragaman variabel-variabel tersebut. Komponen utama yang didapatkan adalah sebanyak k komponen dari sejumlah p variabel yang ada dimana $k \leq p$. Johnson dan Wichern menjelaskan bahwa analisis komponen utama merupakan teknik statistik yang dapat mengubah variabel asli yang saling berkorelasi menjadi satu set variabel baru yang saling bebas (tidak berkorelasi signifikan). Analisis komponen utama dapat mereduksi dimensi data menjadi lebih kecil untuk memudahkan dalam melakukan interpretasi tanpa adanya informasi data yang hilang.

Menurut *Johnson*, secara aljabar komponen utama merupakan teknik rotasi sumbu, sedemikian sehingga sumbu yang baru menjelaskan keragaman yang maksimum. Jika random vektor $X' = (X_1, X_2, \dots, X_p)$ mempunyai matriks kovarians Σ dengan *eigenvalue* $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$.

Komponen utama (Y) dapat dinyatakan sebagai kombinasi linier variabel asal (X) seperti di bawah ini :

$$Y_1 = a_1'X = a_{11}X_1 + a_{21}X_2 + \dots + a_{p1}X_p$$

$$Y_2 = a_2'X = a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p$$

•
•
•
•

$$Y_l = a_l'X = a_{l1}X_1 + a_{l2}X_2 + \dots + a_{lp}X_p$$

Atau dapat juga ditulis dengan notasi $Y_i = a_i'X$, dimana a adalah matriks transformasi yang akan mengubah variabel asal X menjadi Y yang disebut komponen utama, dengan:

$Var(Y_i) = a_i' \Sigma a_i$; $i = 1, 2, \dots, p$ dan $Cov(Y_i, Y_k) = a_i' \Sigma a_k$; $i, k = 1, 2, \dots, p$. Komponen utama tersebut adalah saling bebas dengan varians yang maksimum. Sedangkan untuk pembentukan komponen utama adalah sebagai berikut:

1) Komponen Utama Pertama

Kombinasi linier $a_1'X$ yang memaksimumkan $var(a_1'X)$ dengan ketentuan $a_1'a_1 = 1$ (koefisien yang dinormalkan), $Cov(Y_i, Y_k) = 0$ dimana $k \neq 1$.

2) Komponen Utama Kedua

Kombinasi linier $a_2'X$ yang memaksimumkan $var(a_2'X)$ dengan ketentuan $a_2'a_2 = 1$ dan $Cov(Y_i, Y_k) = 0$ dimana $k \neq 2$.

3) Komponen Utama Ke-i

Kombinasi linier $a_i'X$ yang memaksimumkan $var(a_i'X)$ dengan ketentuan $a_i'a_i = 1$ dan $Cov(a_1'X, a_k'X) = 0$ dimana $k \neq i$, dimana $var(a_i'X)$ mencapai maksimum pada *eigenvalue* λ_1 , atau dapat ditulis:

$$Var(Y_i) = e_i' \Sigma e_i = \lambda_1; i = 1, 2, \dots, p$$

$$Cov(Y_i, Y_k) = e_i' \Sigma e_k; i \neq k$$

Misal $X' = (X_1, X_2, \dots, X_p)$ mempunyai matriks kovarians Σ , dengan *eigenvalue-eigenvektor* $(\lambda_1, e_1), (\lambda_2, e_2), \dots, (\lambda_p, e_p)$, dimana $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$. Jika $Y_1 = e_1'X, Y_2 = e_2'X, \dots, Y_p = e_p'X$ adalah komponen utama, maka:

$$\sigma_{11} + \sigma_{22} + \dots + \sigma_{pp} = (X_i) = \lambda_1 + \lambda_2 + \dots + \lambda_p = (Y_i) = \text{tr}(\Sigma)$$

Merupakan jumlah varians populasi keseluruhan. Sehingga proporsi keseluruhan varians yang dapat diterangkan oleh komponen ke-k adalah sebagai berikut:

$$(\lambda_k) / (\lambda_1 + \lambda_2 + \dots + \lambda_p)$$

dimana $k = 1, 2, \dots, p$.

Koefisien korelasi antara komponen Y_i dengan variabel asal X_k adalah $\rho_{Y_i X_k} =$ untuk $i, k = 1, 2, \dots, p$.

Jumlah komponen utama dapat ditentukan berdasarkan proporsi keragaman data yang dijelaskan oleh komponen utama tersebut. Kaiser dalam buku *Multivariate Analysis Methods and Application* berpendapat

bahwa penentuan jumlah komponen utama didasarkan pada nilai akar karakteristik (*eigen value*) yang lebih dari satu. Untuk mengetahui apakah metode sampling yang digunakan memenuhi syarat atau tidak maka dilakukan uji KMO. Uji ini juga digunakan untuk mengetahui apakah data yang dimiliki bisa dianalisis lebih lanjut dengan analisis faktor atau tidak. Uji KMO dihitung dengan formula sebagai berikut:

$$KMO = \frac{\sum_i \sum_{j \neq i} r_{ij}^2}{\sum_i \sum_{j \neq i} r_{ij}^2 + \sum_i \sum_{j \neq i} a_{ij}^2} ; i = 1, 2, \dots, p ; j = 1, 2, \dots, p$$

Dimana: r_{ij} = Koefisien korelasi sederhana antara peubah i dan j

a_{ij} = Koefisien korelasi parsial antara peubah i dan j

Penilaian uji KMO dari matrik antar peubah adalah sebagai berikut:

1. $0,9 \leq KMO \leq 1,00$ = data sangat baik untuk analisis faktor
2. $0,8 \leq KMO \leq 0,9$ = data baik untuk analisis faktor
3. $0,7 \leq KMO \leq 0,8$ = data agak baik untuk analisis faktor
4. $0,6 \leq KMO \leq 0,7$ = data lebih dari cukup untuk analisis faktor
5. $0,5 \leq KMO \leq 0,6$ = data cukup untuk analisis faktor
6. $KMO \leq 0,5$ = data tidak layak untuk analisis faktor

Selain itu pendekatan lain dinamakan dengan “*Scree Test*”. Dalam pendekatan ini akar karakteristik dari masing-masing komponen utama diplot berdasarkan urutan besarnya akar karakteristik. Dasar pemikirannya adalah bahwa hasil komponen utama sudah diurutkan berdasarkan besarnya akar karakteristik, sehingga komponen utama pertama akan muncul pertama, dan kemudian diikuti oleh komponen utama lainnya yang mempunyai proporsi keragaman yang lebih kecil.

Sebelum melakukan analisis komponen utama terlebih dahulu dilakukan pengecekan apakah data berdistribusi normal multivariat dan

matriks korelasinya bukan merupakan matriks identitas. Pengecekan data berdistribusi normal multivariat dapat dilakukan dengan cara membuat *q-q plot* dari nilai jarak kuadrat pada setiap pengamatan, yaitu $d^2_j = (\mathbf{X}_j - \bar{\mathbf{X}})' \mathbf{S}^{-1}(\mathbf{X}_j - \bar{\mathbf{X}})$; $j = 1, 2, \dots, n$ dan membandingkannya dengan $\chi^2_{p, 0.5}$.

Tahapan pembuatan *q-q plot* adalah sebagai berikut:

1. Menentukan nilai vektor rata-rata
2. Menentukan nilai matriks varians kovarians \mathbf{S}
3. Menentukan nilai jarak *Mahalanobis* setiap titik pengamatan dengan vektor rata-ratanya $d^2_j = (\mathbf{X}_j - \bar{\mathbf{X}})' \mathbf{S}^{-1}(\mathbf{X}_j - \bar{\mathbf{X}})$, $j = 1, 2, \dots, n$
4. Mengurutkan nilai d^2_j dari kecil ke besar, $d^2(1) \leq d^2(2) \leq \dots \leq d^2(n)$
5. Menentukan nilai $p_j = \frac{j}{n+1}$, $j = 1, 2, \dots, n$
6. Menentukan nilai $q_j(p_j)$ sedemikian sehingga $q_j(p_j) = \chi^2_{p, (n-j+1/2)/n}$
7. Membuat *scatter plot* $d^2(j)$ dengan $q_j(p_j)$

Jika *scatter plot* cenderung membentuk garis lurus dan lebih dari 50 % nilai $d^2_j \leq \chi^2_{p, 0.5}$ maka data berdistribusi normal multivariat.

Analisis Cluster

Analisis *cluster* adalah teknik analisis statistika yang digunakan untuk mengidentifikasi dan mengelompokkan objek yang serupa atau memiliki kesamaan ciri dengan memperhatikan kriteria tertentu. Analisis ini merupakan bagian dari analisis statistik multivariat interdependensi yang berfungsi untuk meringkas data dengan cara mengelompokkan objek-objek berdasarkan karakteristik tertentu.

Jika terdapat n objek dan p variat, maka observasi x_{ij} dengan $i = 1, 2, \dots, n$ dan $j = 1, 2, \dots, p$, dapat digambarkan sebagai berikut:

$$KMO = \frac{\sum_i \sum_{j \neq i} r_{ij}^2}{\sum_i \sum_{j \neq i} r_{ij}^2 + \sum_i \sum_{j \neq i} d_{ij}^2}; i = 1, 2, \dots, p; j = 1, 2, \dots, p$$

Adapun ciri-ciri *cluster* adalah adanya homogenitas (kesamaan) yang tinggi antar anggota dalam satu *cluster* (*within-cluster*) dan heterogenitas (perbedaan) yang tinggi antar *cluster* yang satu dengan *cluster* yang lainnya (*between-cluster*). Proses dari analisis *cluster* adalah sebagai berikut:

Menentukan ukuran ketakmiripan antar dua objek

Ukuran ketakmiripan dihitung berdasarkan fungsi jarak antara objek a dan b yang dinotasikan dengan $d(a,b)$ dengan sifat $d(a,b) \geq 0$, $d(a,a) = 0$, $d(a,c) \leq d(a,b) + d(b,c)$, dan $d(a,b)$ meningkat seiring semakin tidak miripnya objek a dan b.

Ukuran jarak yang sering digunakan adalah jarak *euclidean* dengan formula:

$$d_{ih} = \sum_{j=1}^p (x_{ij} - x_{hj})^2$$

dimana:

d_{ih} = jarak kuadrat Euclidean antar objek ke-i dengan objek ke-h.

p = jumlah variabel cluster.

X_{ij} = nilai atau data dari objek ke-i pada variabel ke-j

X_{hj} = nilai atau data dari objek ke-h pada variabel ke-j (Everitt, 1993).

Membuat Cluster dengan Metode Hirarki (Hierarchial Cluster)

Metode ini dimulai dengan mengelompokkan dua atau lebih objek

yang memiliki kesamaan paling dekat. Langkah tersebut dilakukan terus sehingga membentuk semacam pohon yang memiliki tingkatan (hirarki) yang jelas antar objek. Metode yang digunakan untuk pengelompokan secara hirarki adalah metode pautan tunggal (*single linkage*). Metode ini akan mengelompokkan dua objek yang memiliki jarak terdekat. Hasil dari *single linkage clustering* disajikan dalam bentuk dendogram atau diagram pohon dimana cabang-cabang pohon menunjukkan kelompok atau *cluster*.

Selain metode *single linkage* dapat pula menggunakan metode pautan rata-rata (*average linkage*). Metode ini hampir sama dengan *single linkage* hanya jarak yang digunakan adalah rata-rata jarak dari semua objek dalam satu kelompok dengan objek lain di luar kelompok tersebut. Secara umum langkah-langkah metode pengelompokan adalah sebagai berikut:

1. Membuat matriks jarak antar individu **D = (dik)**
2. Mencari nilai jarak terdekat antara dua individu
3. Menghitung jarak antara kelompok dengan beberapa individu di luar kelompok dengan menggunakan metode pautan, misal (**UV**) dengan **W**
4. Menyusun kembali matriks jarak yang baru dengan cara membuat baris dan kolom baru dari nilai jarak antara kelompok dengan beberapa individu di luar kelompok atau kelompok sisa
5. Mengulangi langkah 2-4 sampai semua individu berada dalam satu kelompok dan tidak dapat dikelompokkan lagi
6. Membuat dendogram
7. Memotong dendogram menggunakan selisih jarak terbesar

Tahap akhir dari analisis kelompok adalah terbentuknya kelompok yang diharapkan mempunyai kesamaan karakteristik. Hasil pengelompokan ini digambarkan dalam diagram pohon atau dendogram yang memperlihatkan kelompok dalam tahap pengelompokan. Penentuan banyaknya kelompok dapat dilihat dari pemotongan dendogram pada selisih jarak penggabungan terbesar atau dengan melihat tingkat kemiripan dalam kelompok.

Indikator Pendidikan

Indikator pendidikan merupakan alat ukur yang digunakan untuk melihat seberapa baik kualitas pendidikan yang sudah ada. Indikator pendidikan menjadi sangat penting mengingat adanya permasalahan yang mendasar dalam pendidikan seperti angka putus sekolah yang masih cukup tinggi, adanya kesenjangan mendapatkan kesempatan pendidikan, serta kualitas pendidikan yang belum bisa memenuhi kebutuhan lapangan kerja yang semakin kompetitif. Indikator pendidikan yang digunakan dalam penelitian ini adalah angka melek huruf, angka partisipasi sekolah, angka partisipasi murni, dan angka partisipasi kasar berdasarkan kelompok umur tertentu.

Angka melek huruf adalah proporsi penduduk yang dapat membaca dan atau menulis huruf Latin atau huruf lainnya terhadap jumlah penduduk. Angka partisipasi sekolah adalah proporsi anak sekolah pada usia jenjang pendidikan tertentu dalam kelompok usia yang sesuai dengan jenjang pendidikan tersebut. Angka partisipasi murni adalah proporsi anak sekolah pada satu kelompok usia tertentu yang bersekolah pada jenjang yang sesuai dengan kelompok usianya. Angka partisipasi kasar adalah proporsi anak sekolah pada suatu jenjang tertentu dalam kelompok usia yang sesuai dengan jenjang pendidikan tersebut.

METODOLOGI PENELITIAN

Sumber Data

Data indikator pendidikan Kabupaten/Kota di Jawa Timur tahun 2013 merupakan data sekunder yang diperoleh dari publikasi Badan Pusat Statistik (BPS) Provinsi Jawa Timur.

Variabel Penelitian

Variabel yang digunakan dalam penelitian ini adalah sebagai berikut:

X1 : angka melek huruf (%)

X2 : angka partisipasi sekolah 7-12 tahun (%)

X3 : angka partisipasi sekolah 13-15 tahun (%)

X4 : angka partisipasi sekolah 16-18 tahun (%)

X5 : angka partisipasi kasar 7-12 tahun (%)

X6 : angka partisipasi kasar 13-15 tahun (%)

X7 : angka partisipasi kasar 16-18 tahun (%)

X8 : angka partisipasi murni 7-12 tahun (%)

X9 : angka partisipasi murni 13-15 tahun (%)

X10: angka partisipasi murni 16-18 tahun (%)

Langkah-langkah dalam analisis data adalah sebagai berikut:

1. Melakukan pengujian korelasi antar variabel
2. Melakukan analisis komponen utama jika terdapat korelasi yang signifikan antar variabel
3. Melakukan analisis *cluster* dengan metode *hierarchial cluster*
4. Melakukan perbandingan nilai rata-rata dari *cluster* yang terbentuk

ANALISIS DAN PEMBAHASAN

1. Analisis Korelasi

Salah satu asumsi yang harus dipenuhi dalam analisis *cluster* adalah antar variabel indikator harus bebas dari korelasi. Untuk melihat ada tidaknya korelasi antar variabel indikator, dilakukan analisis korelasi dengan hipotesis sebagai berikut:

H₀: Tidak terdapat korelasi antara variabel indikator pendidikan Jawa Timur tahun 2013

H₁ : Terdapat korelasi antara variabel indikator pendidikan Jawa Timur tahun 2013

Hasil korelasi anatar variabel indikator adalah sebagai berikut:

	ANIH	APS_7_12	APS_1	APS_1	APK_7_12	APK_1	APK_1	APAL_7_12	APAL_1	APAL_16_18
ANIH										
Pearson Correlation	1	.209	.315	.618	-.310 ^{**}	.236	.619 ^{**}	.080	.439 ^{**}	.589 ^{**}
Sig. (2-tailed)		.208	.006	.000	.001	.154	.006	.635	.006	.000
N	36	36	36	36	36	36	36	36	36	36
APS_7_12										
Pearson Correlation	.209	1	.285	.275	-.315	.427 ^{**}	.266	-.070	.326 ^{**}	.335 ^{**}
Sig. (2-tailed)	.208		.083	.094	.059	.007	.107	.676	.046	.040
N	36	36	36	36	36	36	36	36	36	36
APS_13_15										
Pearson Correlation	.540 ^{**}	.285	1	.693 ^{**}	-.534 ^{**}	.572 ^{**}	.672 ^{**}	-.171	.644 ^{**}	.357 ^{**}
Sig. (2-tailed)	.000	.083		.000	.000	.000	.000	.001	.000	.000
N	36	36	36	36	36	36	36	36	36	36
APS_16_18										
Pearson Correlation	.347 ^{**}	.275	.653 ^{**}	1	-.490 ^{**}	.618 ^{**}	.839 ^{**}	-.064	.633 ^{**}	.389 ^{**}
Sig. (2-tailed)	.000	.094	.000		.002	.000	.000	.000	.000	.000
N	36	36	36	36	36	36	36	36	36	36
APK_7_12										
Pearson Correlation	-.310 ^{**}	-.313	-.534 ^{**}	-.490 ^{**}	1	-.693 ^{**}	-.542 ^{**}	.499 ^{**}	-.531 ^{**}	.533 ^{**}
Sig. (2-tailed)	.001	.055	.001	.002		.000	.000	.000	.000	.001
N	36	36	36	36	36	36	36	36	36	36
APK_13_15										
Pearson Correlation	.236	.427 ^{**}	.375 ^{**}	.618 ^{**}	-.693 ^{**}	1	.480 ^{**}	-.503 ^{**}	.809 ^{**}	.621 ^{**}
Sig. (2-tailed)	.154	.007	.000	.000	.000		.013	.001	.000	.000
N	36	36	36	36	36	36	36	36	36	36
APK_16_18										
Pearson Correlation	.619 ^{**}	.266	.672 ^{**}	.839 ^{**}	-.542 ^{**}	.400 ^{**}	1	-.139	.438 ^{**}	.862 ^{**}
Sig. (2-tailed)	.000	.107	.000	.000	.013	.000		.405	.006	.000
N	36	36	36	36	36	36	36	36	36	36
APAL_7_12										
Pearson Correlation	.080	.676	-.171	.703	.499 ^{**}	-.503 ^{**}	.405	1	-.007	.642
Sig. (2-tailed)	.635	.001	.001	.001	.001	.001	.001		.965	.000
N	36	36	36	36	36	36	36	36	36	36
APAL_13_15										
Pearson Correlation	.439 ^{**}	.326 ^{**}	.644 ^{**}	.633 ^{**}	-.531 ^{**}	.809 ^{**}	.438 ^{**}	-.007	1	.720 ^{**}
Sig. (2-tailed)	.006	.046	.000	.000	.000	.000	.006	.000		.000
N	36	36	36	36	36	36	36	36	36	36
APAL_16_18										
Pearson Correlation	.589 ^{**}	.335 ^{**}	.663 ^{**}	.941 ^{**}	-.533 ^{**}	.621 ^{**}	.663 ^{**}	.642	.720 ^{**}	1
Sig. (2-tailed)	.000	.040	.000	.000	.001	.000	.000	.000	.000	
N	36	36	36	36	36	36	36	36	36	36

**. Correlation is significant at the 0.01 level (2-tailed).

Tabel 1 Nilai Korelasi Antar Variabel Indikator Pendidikan

Dari tabel di atas dapat disimpulkan bahwa terdapat korelasi antar variabel indikator pendidikan Jawa Timur tahun 2013. Hal ini dikarenakan nilai , pada taraf nyata 5% sehingga H_0 ditolak. Selanjutnya dilakukan analisis komponen utama untuk menghasilkan skor komponen utama yang saling bebas dari korelasi.

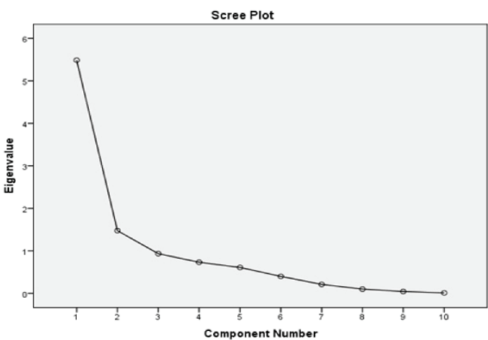
2. Analisis Komponen Utama

Tabel 2 Nilai Kaiser-Meyer-Olkin (KMO)

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.578
Bartlett's Test of Sphericity	Approx. Chi-Square	367.680
	df	45
	Sig.	.000

Hasil dari analisis komponen utama menunjukkan bahwa nilai KMO = 0.578 yang dapat disimpulkan bahwa data cukup untuk dilakukan analisis faktor.

Hasil analisis faktor diperoleh dua *eigen value* yang lebih dari 1 yang menunjukkan bahwa didapatkan dua faktor yang mewakili 10 indikator pendidikan di Jawa Timur tahun 2013 seperti terlihat dari *scree plot* di bawah ini.:

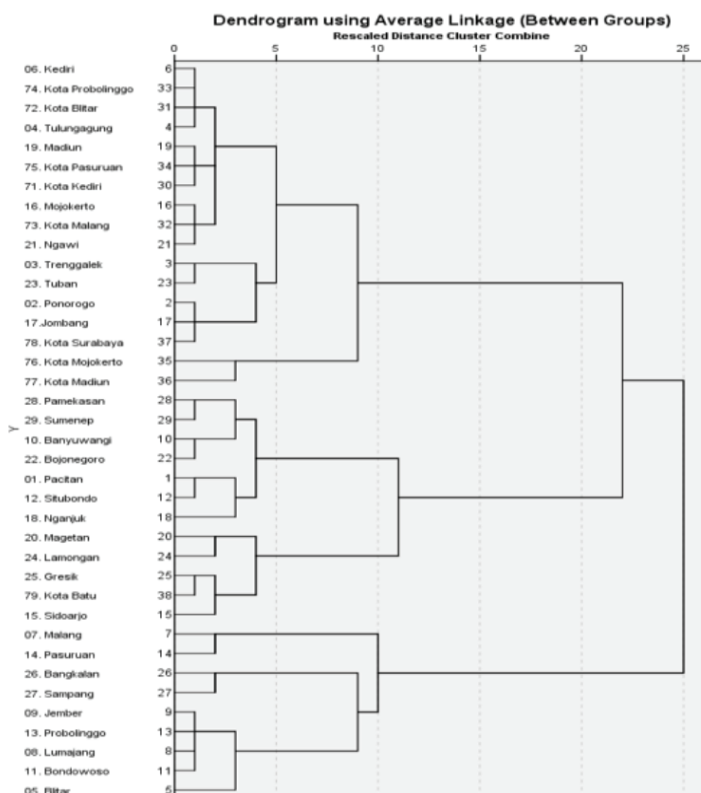


Gambar 1 *Scree Plot* dari Hasil Analisis Komponen Utama

Selanjutnya dilakukan analisis *hierarchial cluster* terhadap dua faktor yang terbentuk.

3. Analisis Hierarchial Cluster

Langkah selanjutnya adalah pengelompokan dengan menggunakan jarak Euclid pautan rata-rata. Dari hasil output SPSS 16 diperoleh matriks *proximity*, dengan angka yang tertera adalah jarak antara dua buah variabel. Seperti diketahui, langkah pertama pada analisis *cluster* adalah melakukan pengukuran terhadap kesamaan (*similarity*) antar variabel, sesuai dengan tujuan *cluster* untuk mengelompokkan variabel yang sama (*similar*).



Gambar 2 Dendrogram Hasil *Cluster* yang Terbentuk

Setelah jarak antar variabel diukur dengan cara *Euclidean*, selanjutnya dilakukan pengelompokan variabel secara hierarkis. Cara hierarki berarti pengelompokan dilakukan secara bertingkat, satu demi satu, atau dari terbentuknya *cluster* yang banyak, selanjutnya jumlah *cluster* berkurang sehingga akhirnya semua menjadi satu *cluster* saja. Cara ini dinamakan agglomerasi. Proses agglomerasi pada akhirnya akan menyatukan semua variabel menjadi satu *cluster*. Hanya dalam prosesnya, dihasilkan beberapa *cluster* dengan masing-masing anggotanya, tergantung jumlah *cluster* yang terbentuk. Perincian *cluster* yang terbentuk dapat dilihat pada dendrogram dalam Gambar 2 di atas.

Penentuan banyak kelompok dapat dilihat dari pemotongan dendrogram pada selisih jarak penggabungan terbesar. Dari pemotongan dendrogram tersebut diperoleh dua kluster. Perincian kluster yang terbentuk adalah sebagai berikut:

Kediri, Kota Probolinggo, Kota Blitar, Tulungagung, Madiun, Kota Pasuruan, Kota Kediri, Mojokerto, Kota Malang, Ngawi, Trenggalek, Tuban, Ponorogo, Jombang, Kota Surabaya, Kota Mojokerto, Kota Madiun, Pamekasan, Sumenep, Banyuwangi, Bojonegoro, Pacitan, Situbondo, Nganjuk, Magetan, Lamongan, Gresik, Kota Batu, dan Sidoarjo. Malang, Pasuruan, Bangkalan, Sampang, Jember, Probolinggo, Lumajang, Bondowoso dan Blitar.

Setelah diperoleh dua kluster kabupaten/kota yang mewakili tingkat pendidikan yang relatif sama, selanjutnya dilakukan perbandingan nilai rata-rata dari kluster yang terbentuk. Berikut adalah rata-rata setiap variabel dari dua kluster yang terbentuk:

Tabel 3 Nilai Rata-rata Variabel Masing-masing *Cluster* yang Terbentuk

Variabel	<i>Cluster</i> 1	<i>Cluster</i> 2
X_1	91.91	84.45
X_2	99.19	98.4
X_3	95.52	86.9
X_4	68.45	51.96
X_5	104.3	109.7

X_6	95.31	79.73
X_7	69.02	50.91
X_8	95.89	97.07
X_9	82.01	68.8
X_{10}	59.37	42.37

Dari tabel 3 di atas dapat dikelompokkan bahwa *cluster* 1 memiliki tingkat pendidikan yang lebih tinggi dibandingkan *cluster* 2. *Cluster* 1 memiliki rata-rata angka melek huruf, angka partisipasi sekolah 7-12 tahun, angka partisipasi sekolah 13-15 tahun, angka partisipasi sekolah 16-18 tahun, angka partisipasi kasar 13-15 tahun, angka partisipasi kasar 16-18 tahun, angka partisipasi murni 13-15 tahun, dan angka partisipasi murni 16-18 tahun yang lebih tinggi daripada *cluster* 2. Dibandingkan dengan *cluster* 1, *cluster* 2 memiliki keunggulan pada nilai rata-rata angka partisipasi kasar dan angka partisipasi murni pada jenjang usia 7-12 tahun.

KESIMPULAN

Berdasarkan hasil analisis *hierarchial cluster* terhadap data indikator pendidikan tiap Kabupaten/Kota di Jawa Timur tahun 2013 dapat disimpulkan bahwa terbentuk 2 *cluster*. *Cluster* pertama terdiri dari 29 kabupaten/kota yaitu, Kediri, Kota Probolinggo, Kota Blitar, Tulungagung, Madiun, Kota Pasuruan, Kota Kediri, Mojokerto, Kota Malang, Ngawi, Trenggalek, Tuban, Ponorogo, Jombang, Kota Surabaya, Kota Mojokerto, Kota Madiun, Pamekasan, Sumenep, Banyuwangi, Bojonegoro, Pacitan, Situbondo, Nganjuk, Magetan, Lamongan, Gresik, Kota Batu, dan Sidoarjo yang memiliki tingkat pendidikan lebih tinggi dibanding *cluster* 2. *Cluster* kedua terdiri dari 9 kabupaten/kota yaitu, Malang, Pasuruan, Bangkalan, Sampang, Jember, Probolinggo, Lumajang, Bondowoso dan Blitar.

DAFTAR PUSTAKA

- Dillon, W.R. and Goldstein, M. *Multivariate Analysis Methods and Application*. New York: John Wiley and Sons, 1984.
- Everitt, B.S. dan Dunn, G. *Cluster Analysis Second Edition*. Heineman London: Education Book, Ltd, 1980.
- Everitt, B.S. *Cluster Analysis*. Third Edition. Halsted Press an Imprint of New York: John Wiley and Sons Inc, 1993.
- Hendri, Davy. *Kaji Ulang Indikator Pembangunan Pendidikan*. http://www.targetmdgs.org/index.php?option=com_content&task=view&id=301&Itemid=6). 2008.
- Johnson, R.A. and Wichern, D.W. *Applied Multivariate Statistical Analysis*. New Jersey: Prentice Hall Inc. 1982.
- Johnson, R.A. and Wichern, D.W. *Applied Multivariate Statistical Analysis*. New Jersey: Prentice Hall Inc., 2002.
- Jolliffe. *Principal Component Analysis*. New York: Springer Verlag, 2002.
- Sahputra, Ade, Peran UNESCO dalam Pemberantasan Buta Aksara di Indonesia tahun 2007-2012, *Jom FISIP*, Volume 1 No. 2-Oktober 2014